ARMY RESEARCH LABORATORY

Network Science for Decision-making: Impact of Distributed Information Quality on Performance of Decision-making Groups

by Kevin Chan and Brian Rivera

ARL-TR-5454 **February 2011**

NOTICES

Disclaimers

Army Research Laboratory

Adelphi, MD 20783-1197

ARL-TR-5454 **February 2011**

Network Science for Decision-making: Impact of Distributed Information Quality on Performance of Decision-making Groups

Kevin Chan and Brian Rivera
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* February 2011	**2. REPORT TYPE** DSI	**3. DATES COVERED (From - To)**
4. TITLE AND SUBTITLE Network Science for Decision-making: Impact of Distributed Information Quality on Performance of Decision-making Groups		**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Kevin Chan and Brian Rivera		**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-CIN-T 2800 Powder Mill Road Adelphi MD 20783-1197		**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-TR-5454
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Within tactical networks, communication networks support command and control organizations and social networks by enabling the flow of information. The connectivity and quality of service of the communication network results in a varying degree of mission performance. We consider this problem from a network science perspective, assuming that these networks are comprised of communication, social/cognitive, and information networks; and are characterized by the interactions between these constituent networks. We previously proposed a general framework of human trust in networks that models trust as a composite of reliability and availability of the network services. This work considers the relationship between individual and team decision making and communication networks and how these interactions affect individual and group performance. Through the use of a command and control experiment platform called the Experimental Laboratory for Investigating Collaboration, Information-sharing, and Trust (ELICIT), we study the impact of loss and delay in a distributed server scenario using human-agent models.

15. SUBJECT TERMS

Network Science, decision making, human agents

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Kevin Chan
a. REPORT UNCLASSIFIED	**b. ABSTRACT** UNCLASSIFIED	**c. THIS PAGE** UNCLASSIFIED	UU	24	**19b. TELEPHONE NUMBER** (*Include area code*) (301) 394-5640

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

1. Introduction/Objective

Tactical networks are complex networks, which are comprised of communication, information and social/cognitive networks. Interactions occur within and between these composite network layers. Presently, there is a lack of understanding of the interaction and dynamics between these networks. The research area of network science seeks to find interactions within complex networks to enable optimal design, prediction, and modeling of tactical networks (*1*). Figure 1 shows the network layers. For example, most current communication network research involves quantifying or modeling performance with regard to technical measures (i.e., bandwidth, capacity, and error rates) (*2*–*6*). Studies tend not to focus on the end user of the network. Thus, these studies are only useful in providing objective measures of network performance. Conversely, social/cognitive network research includes studies that concern the end user, but communication models are oversimplified or difficult to observe (*7*–*12*). In tactical environments, we are interested in a cross-layer understanding so that we can model and identify tradeoffs between performance and resource consumption within each of the network layers. Specifically, we are interested in determining how communication network parameters affect human performance. We address the impact of loss and delay in communications on decision making and situational awareness in a specific experimental scenario.

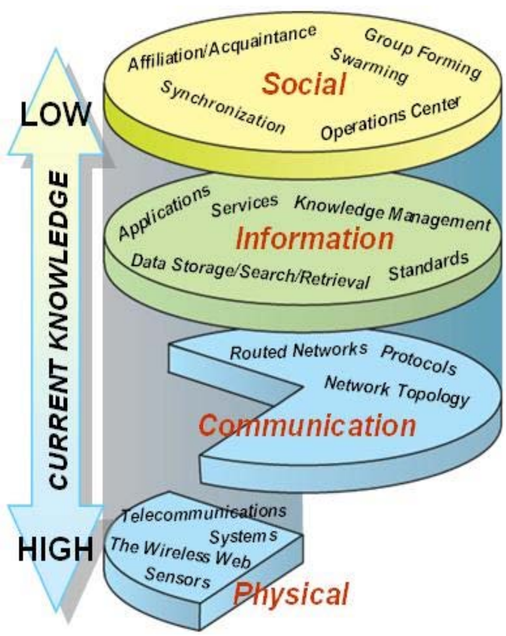

Figure 1. Network science network layers.

The goal of this research effort is to investigate the impact of distributed information quality on the performance of decision-making groups in networks. Specifically, we are interested in examining the performance of individuals and teams in a networked environment in a distributed

server scenario. As shown in figure 2, this scenario consists of information sources, a set of distributed information servers, and a group of decision-makers. Information sources will generate reports of events to populate the distributed information servers. Decision-makers will periodically access these information servers to gain information and establish situational awareness to execute their mission. Depending on the information quality and flow of information, the decision-makers will perform accordingly.

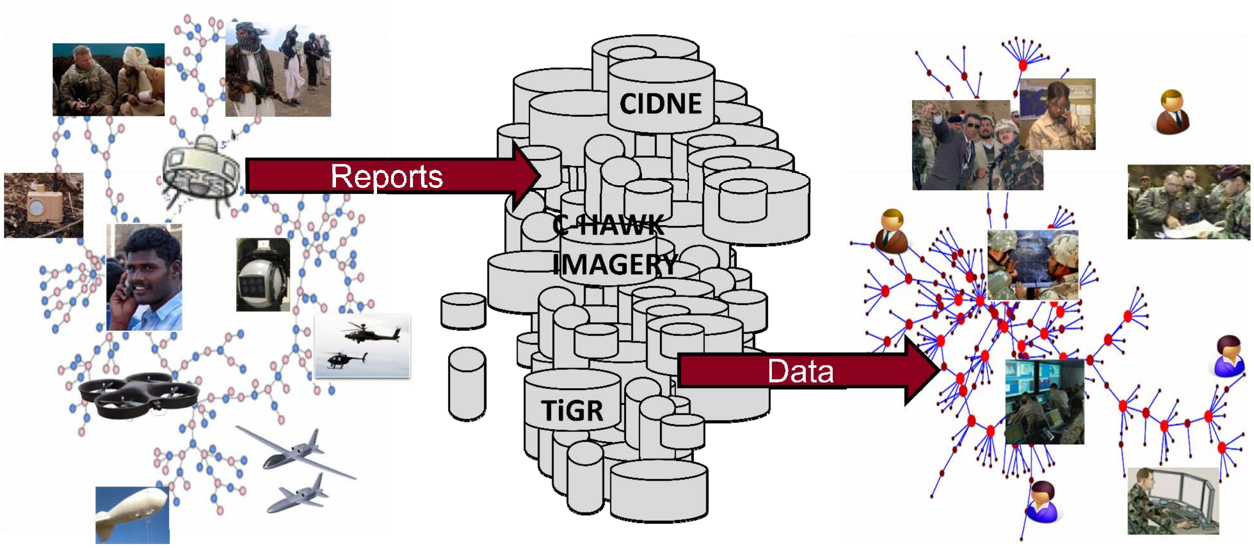

Figure 2. Illustration of distributed server scenario and information flow.

The arrangement of the servers and the performance of the underlying communication network causes the information to be disseminated at various rates and results in periods where the servers are not consistent with each other. This variation in the quality of information that users receive when requesting information from the servers affects decision making, mission performance, and ultimately, trust in the network. We consider the accessibility, accuracy, and freshness of the information retrieved from queries on the servers. Accessibility pertains to the delay or availability of the servers, and accuracy implies the synchronization with the other servers in the network. Freshness includes the quality (timeliness) of the information in the servers with respect to the current situation.

2. Approach

The Experimental Laboratory for Investigating Collaboration, Information-sharing, and Trust (ELICIT), a command and control (C2) experiment platform, has been used to conduct these experiments (*13*). ELICIT is a configurable software platform designed to measure the behavior of social networks in a command and control information-sharing scenario. Participants in ELICIT experiments are periodically provided with “factoids” or snippets of information. These

factoids are sent and received among the participants, or the participants can retrieve information from a set of simulated Web sites or databases. This information is used to deduce specific information of a fictional terrorist threat (determining the WHO/WHAT/WHERE/WHEN of the threat). ELICIT is designed to study the organization of social networks and the interactions within these networks. The structure of ELICIT is shown in figure 3. The connectivity of the participants and the Web sites has been specified to fit the scenarios of interest.

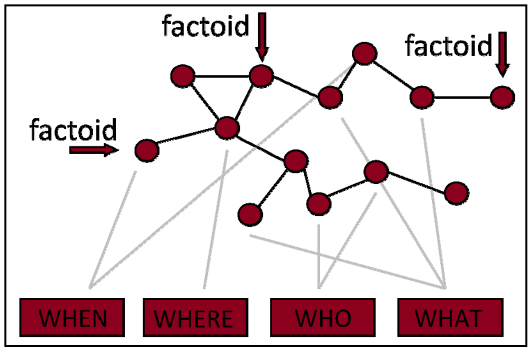

Figure 3. Basic ELICIT organizational structure.

ELICIT aims to demonstrate though experimentation the effectiveness of variations in organizations, as suggested by Alberts (*14*), where it is hypothesized that edge (or flat topologies) organizations will outperform hierarchical organizations in command and control scenarios with uncertainty, complexity, and requirements for agility. ELICIT allows for human-in-the-loop or agent-based experimentation with various organizations (including edge and hierarchical organizations). The underlying communication infrastructure is assumed but not emphasized.

To measure the performance of ELICIT, we evaluate the ability of the group to correctly determine the details of the terrorist threat. The correctness measure represents a measure of situational awareness within this scenario. Correctness is measured by the accuracy of the WHO, WHAT, WHERE, and WHEN. The details WHO, WHAT, WHERE are scored with 0 or 1, and WHEN has a score of {0, 0.25, 0.5, 0.75, and 1.0}, allowing for partial correctness. The overall correctness score, *C*, a value between 0 and 1 is

$$C = 0.25\ (\ \mathrm{WHO} + \mathrm{WHERE} + \mathrm{WHAT} + \mathrm{WHEN}\ ). \tag{1}$$

Recently, sensemaking agents were developed to run ELICIT (*15*, *16*), which enable trials of ELICIT to be run without human participants. These sensemaking agents are governed by a set of parameters, which can be used to characterize the way the agent processes and shares information with other participants in the experiment. Agents and humans are able to participate in the experiments together; however, we only consider experiments comprised solely of agents. It is assumed that the sensemaking agents are valid models of humans in these experiments. Figure 4 shows the sensemaking agent factoid processing steps in terms of mental processing and sharing decisions. A validation of the ability of agents to model actual human behaviors was

performed by Wynn (*16*). When considering communication networks within ELICIT, it is hypothesized that the sensemaking agent parameters can represent communication network parameters.

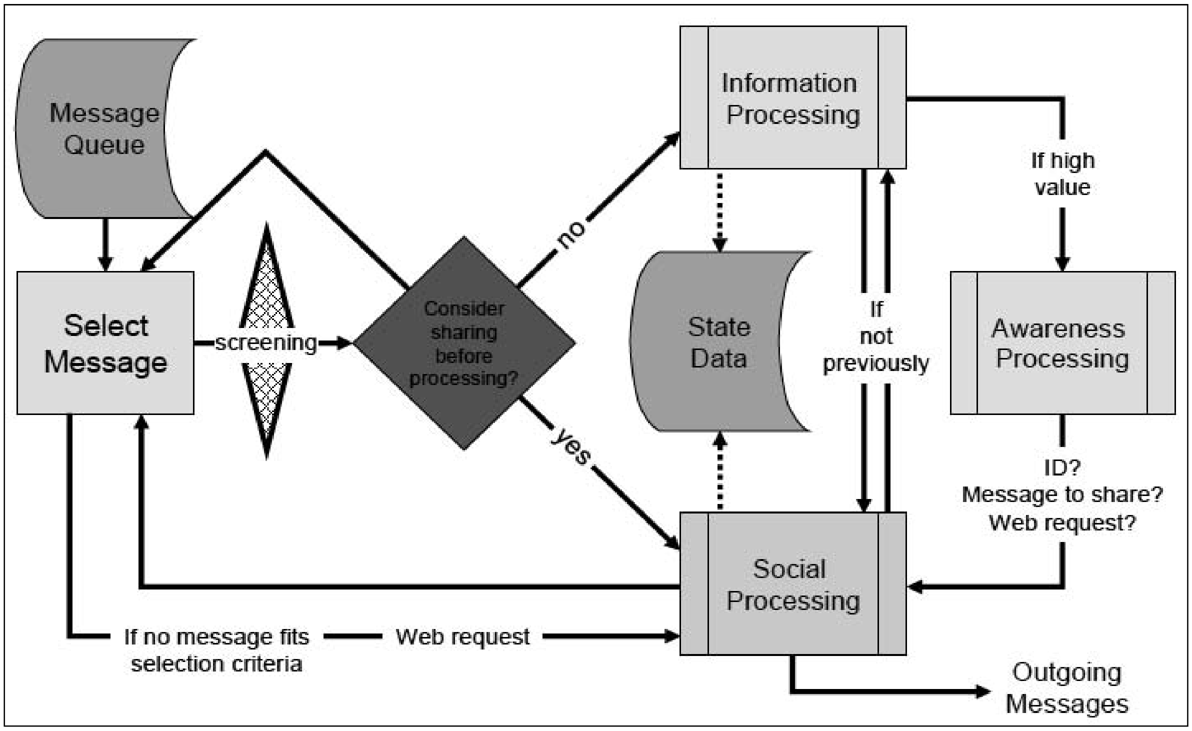

Figure 4. Process flow diagram of ELICIT sensemaking agent model.

In terms of the flow of information, factoids represent intelligence reports, which contain partial information regarding the possible insurgent threat. Initially, the factoids are randomly distributed to the information nodes at the start of the trial (factoids can be distributed in waves over time). As time progresses, the factoids are shared to neighboring nodes and posted to the Web sites, making this information available to the decision-making nodes. The factoid set (there are 68 unique factoids) contains all the information needed to uniquely determine the details of the potential insurgent threat. The factoid set does not contain any conflicting information that would lead to incorrect conclusions. Any errors that occur are the result of errors in the processing of the information.

3. Results

We have run several different sets of experiments with ELICIT to examine the impact of communication networks on the decision-making ability of the organization. Figure 5 shows an adaptation of the original ELICIT structure to represent the distributed server scenario. In this figure, a set of nodes in the communication network represent the information nodes. These

nodes are randomly distributed into a unit square and have communication radius *r*. In this square of side lengths equal to 1, any two nodes within distance *r* can communicate with each other. Additionally, the distributed servers or databases are placed at each of the corners of the unit square (i.e., at (0, 0), (1, 0), (1, 1), (0, 1)). Any node that is within distance *r* of the databases can communicate (thus, post information) to the database. Each of the decision-making nodes is connected to each of the servers, and the nodes experience random loss and delay when communicating with the databases. In terms of the distribution of the factoids, the factoids are randomly distributed to each of the nodes without duplication.

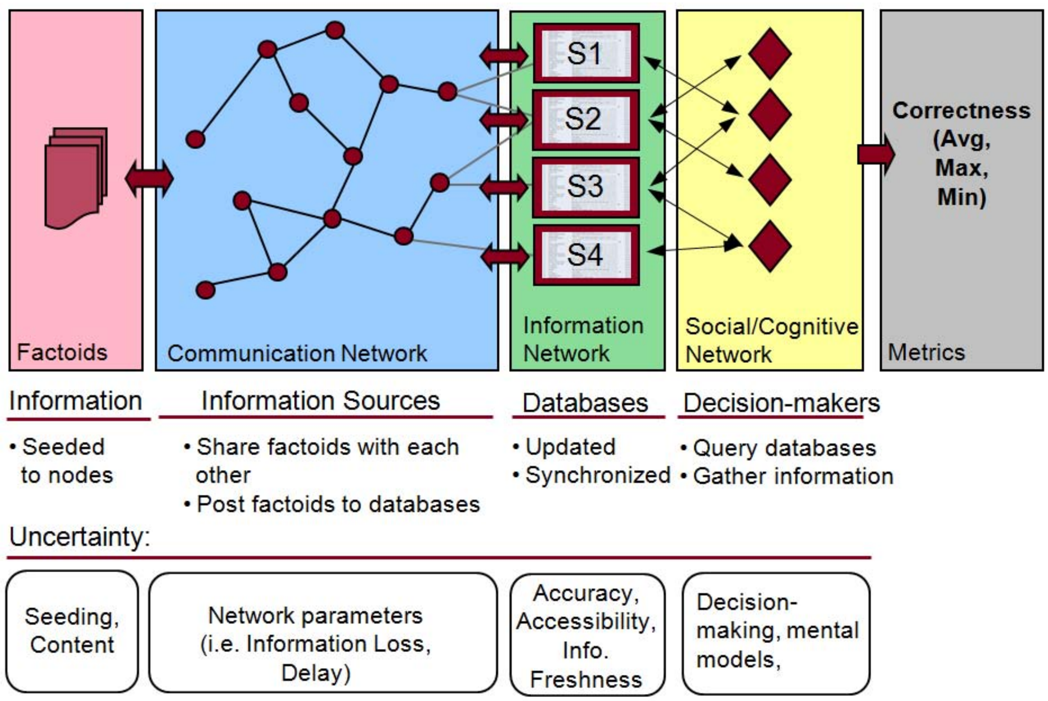

Figure 5. Adapted ELICIT structure for distributed server scenario.

First, we show previously acquired results which study the information flow within the communication network (the set of information sources in figure 5). Loss and delay between the intermediate hops within this network are studied. We measure the awareness of the information nodes in terms of which factoids each of the information source nodes have received. This provides an initial indication of the rate at which information arrives at the servers. Second, we conducted a set of the distributed server scenario experiments, where we measured the performance of a set of decision-making nodes accessing the distributed servers.

3.1 Distributed Networking Scenario

In previous work, the sensemaking agent parameters were used to simulate parameters in the communication network (i.e., loss and delay) (*17*). A distributed networking scenario is used, where a random organization of agents performed the ELICIT task and the decision-making performance of the entire organization was examined (figure 6). In these experiments, only the

information nodes are used, not the distributed information servers, which makes the experiment completely distributed. Each node is required to perform the decision-making task, where in the distributed server scenario, only the performance of decision-making nodes is measured. Overall average correctness is studied as a function of loss and delay in the links during these experiments.

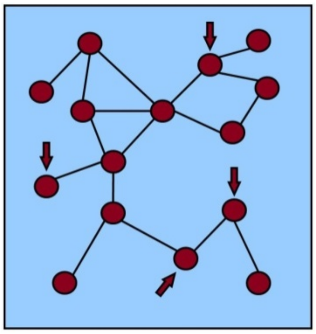

Figure 6. Distributed organization scenario.

These ELICIT trials use a network of 68 nodes, with the communication radius chosen to guarantee connectivity within the organization ($r = 0.3$). The nodes are deployed randomly into a unit area. These trials are 2 h in duration, and the correctness after 1 h, $C(1)$, and 2 h, $C(2)$, is measured. We show that the baseline (8-s packet latency, 0% packet loss) achieves full average correctness after approximately 2 h. By incorporating packet delay and loss in the communications, the average correctness is degraded. The results of 25 Monte Carlo simulations are shown, and the average correctness is measured by averaging across every node in the organization and also averaging across network realizations. A unique network topology is used for each Monte Carlo simulation.

In terms of correctness, we consider the average correctness of all nodes over a set of Monte Carlo simulations. This provides insight into the overall decision-making performance of the group. However, it may be of interest to look at the distribution of the correctness of the individual nodes within the experiments. For example, figure 7 shows the distribution of the nodes who have correctly identified the WHO, WHAT, WHERE, or WHEN for one experiment. This histogram shows the distribution of the number of details correctly identified by the nodes, where 4 indicates that the node has correctly identified all of the details of the potential threat (full correctness). The distribution of the partial correctness of the nodes is shown for after 1 and 2 h for a baseline experiment (0% loss, 8-s delay) for 68 nodes. After 1 h, there appears to be an even distribution across each of the five possible correctness states, and after an additional hour, most of the nodes have achieved full correctness.

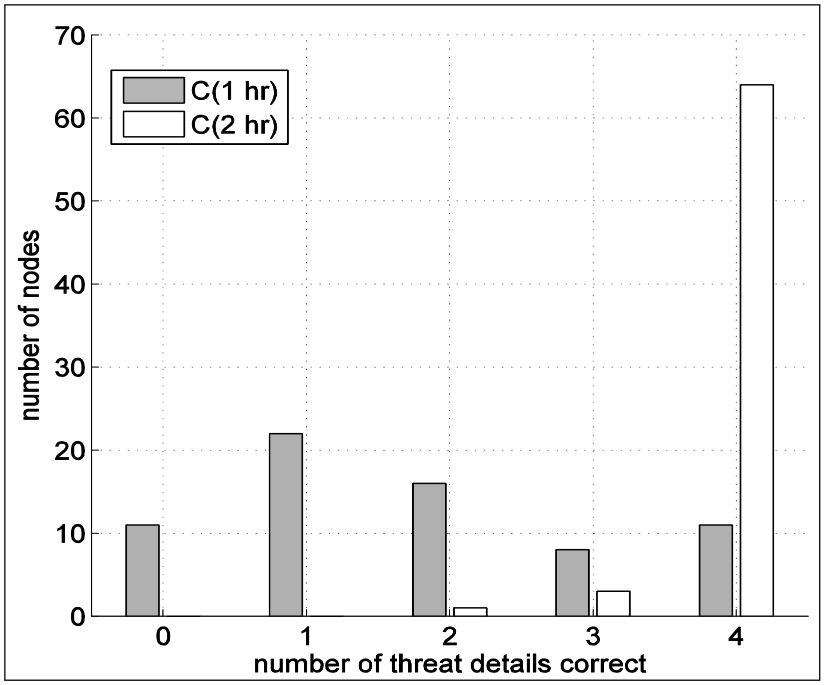

Figure 7. Distribution of number of threat details correct for C(1 h) and C(2 h) for one trial.

Figure 8 shows the results of varying the packet delay within the information source network. Figure 8 is the average correctness of an organization for a set of Monte Carlo simulations. Each trial uses a unique random topology of nodes. The packet latency is varied from 8 s to 5 min. In each trial, the same packet delay is used for every factoid shared from node to node. Figure 8a shows the average correctness versus time for the set of packet latencies. In these trials, there is no packet loss. This figure shows the impact of packet delay on the average correctness. Figure 8b shows the average $C(1)$ and $C(2)$ versus packet delay. This set of experiments demonstrates that the average correctness exhibits a threshold effect for this particular scenario at packet latency of approximately 30 s, as seen in figure 8b. The error bars indicate one standard deviation of the results obtained through experiments; this does indicate a significant variance in the correctness.

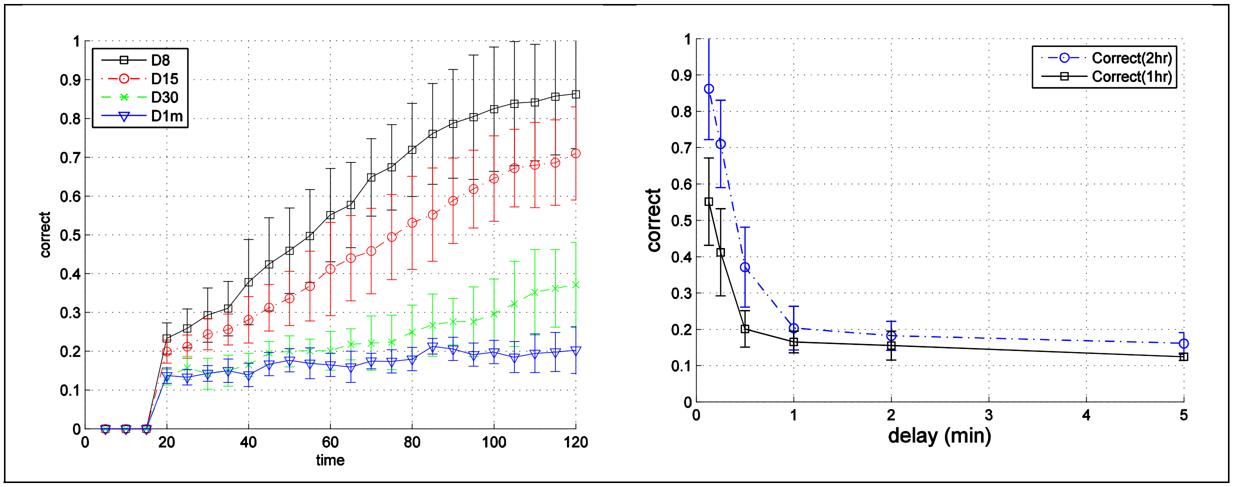

Figure 8. (a) Correctness vs. time for packet loss = 0% and (b) correctness vs. delay for packet loss = 0%.

We also consider the effect of packet loss within the information source network. Packet loss rates 0%, 20%, 40%, 60%, 80%, and 90% are used. In each trial, every factoid is successfully transmitted with probability according to the packet loss rate. Figure 9a is a plot of average correctness versus time for a single trial of each set of packet losses. This shows the degradation in performance as packet loss increases. Figure 9b shows the average C(1 h) and C(2 h) versus time for a set of trials, where each plot line represents a run for a particular packet loss. In these trials, the packet delay is 8 s. The error bars reflect one standard deviation on the gathered results. This figure shows the impact of packet delay on the average correctness. There is a threshold effect for this particular scenario at around a packet loss rate of 50% as seen in figure 9b, but average correctness degrades more gracefully than the performance with regard to delay.

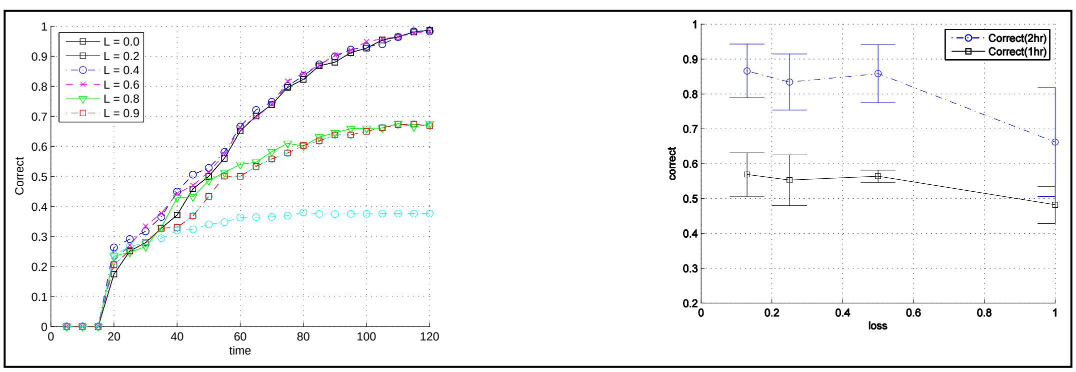

Figure 9. (a) Correctness vs. time for packet delay = 8 s and (b) correctness vs. loss for packet latency = 8 s.

In another set of experiments, both packet loss and packet delay are jointly considered. To illustrate the effect of these two network parameters on correctness within ELICIT trials, we consider each pair of packet loss rates (0%, 20%, 40%, 60%, 80%, and 90%) and packet delays (8 s, 15 s, 30 s, 1 min, 2 min, and 5 min). The same approach to the organization and use of factoids as the previous set of trials is implemented. Figure 10 shows C(2 h) as a function of both

packet loss and packet delay. The average correctness of the network demonstrates a tolerance of packet loss in low packet latency situations. Once the packet latency is greater than around 30 s (as shown in figure 8), the performance of ELICIT drastically decreases, regardless of the packet loss rate. This suggests that the performance is more sensitive to packet delays than loss. The organization is able to handle packet losses with redundancy in the network, whereas with packet delays, the nodes simply have to wait.

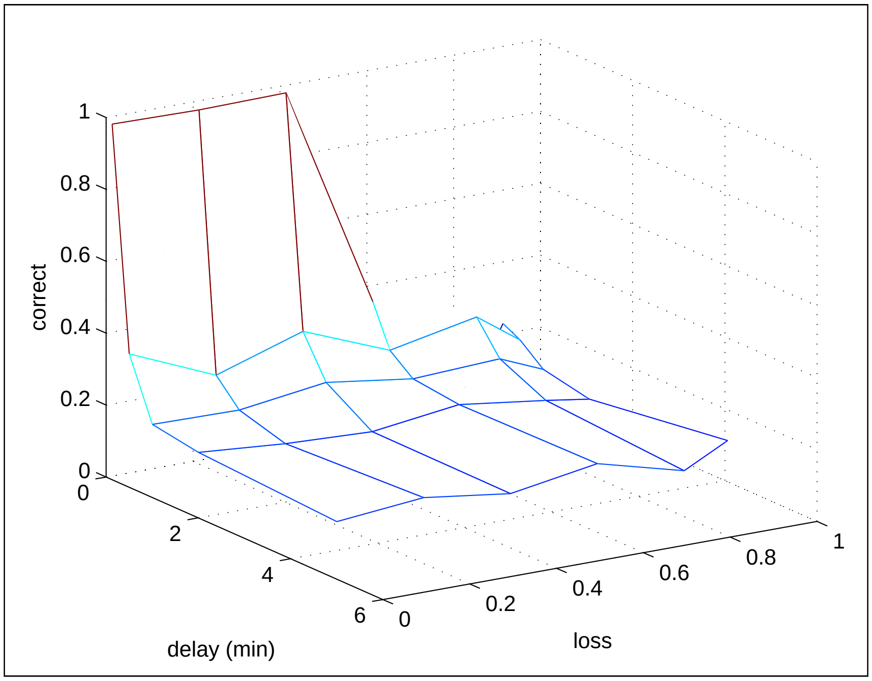

Figure 10. Surface plot of correctness after 2 h for packet delay and loss.

We also investigate the effect of the performance of the organization when connectivity was varied. To vary the connectivity within the organization, the communication radius *r* is varied when creating *G(n,r)*. In terms of the communication network, the cost of increased connectivity is increased energy consumption. Figure 11 shows the performance of the organization versus the communication radius of the organization.

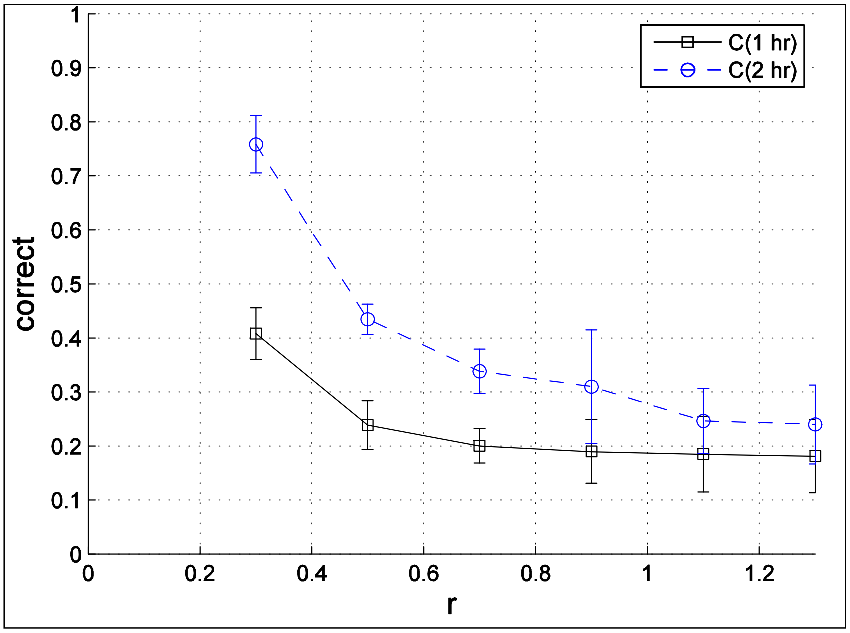

Figure 11. Correctness vs. communication radius.

The result of this set of ELICIT experiments contradicts the expected outcome. We found that with increased connectivity, the average performance of the organization decreases. One would expect that with greater connectivity in the network the performance of the network would improve. By allowing for the nodes to share factoids with nodes farther away, this requires less multi-hop communications so factoids will be disseminated more rapidly.

The immediate explanation is that the agents are suffering from "information overload." Due to the increasing number of neighboring agents in the organization, the nodes are receiving more factoids than the agent is able to process. This is shown in figure 12. The total shares received in each 5-min interval over one realization of a 2-h trial length are plotted for each of the communication radii used to create the topology of the organization. In the regions where $r > 1.0$, the average total number of shares received is saturated, where the nodes in the network are receiving an average of 4000 shares in the course of the 2-h trial. When examining the performance of the network in figure 11 and considering the share behavior in figure 12, this indicates that the nodes are receiving an average of 58 factoids in a 5-min interval, but their performance does not correspond to the number of received factoids. This indicates that the nodes are flooded with factoids and cannot process all of the factoids. The phenomenon of overwhelmed agents explains the observed behavior. So, if the information nodes are exchanging information with a great number of other nodes, then information overload is a possible outcome, which will reduce information flow, and thus, overall correctness.

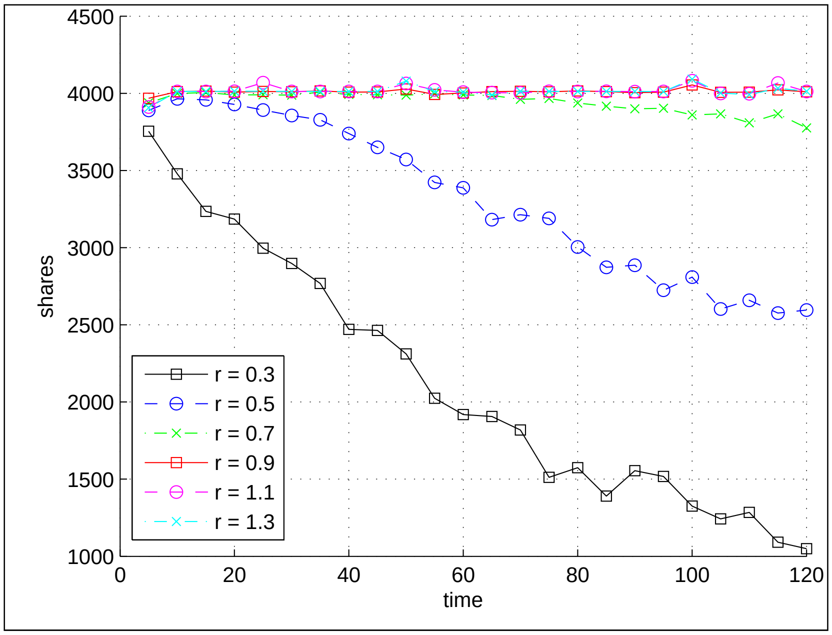

Figure 12. Total shares received in each 5-min interval for various *r*.

3.2 Distributed Server Scenario

We consider the distributed server scenario experiments with ELICIT, as shown in figure 13, by making use of the distributed databases in addition to the set of information nodes. In this case, we studied the impact of communication network performance on the decision-making ability of decision-making nodes in the distributed server scenario. Each of the decision-making nodes are connected to the distributed information servers, but experience various quality of service from each of the servers. Experiments use agents in both the information nodes as well as the decision-making nodes. This results reflect a Monte Carlo simulation of a set of 50 ELICIT trials, each 2 h in duration with a different network topology of 68 information nodes for each trial. The communication radius for the information network is $r = 0.3$, which is sufficient to guarantee connectivity of the information node network and the distributed servers. In the communications between the information nodes, each transmission experiences 0% loss and 8-s delay. There are four distributed servers, which provide a different combination of information loss and information delay in server access attempts of each decision-making agent. These servers are placed in the corners of the unit area of the information network. Information loss $L = \{ 0, 0.2, 0.4, 0.6, 0.8, 0.9 \}$ represents an unsuccessful server access attempt. For the parameter information delays, $D = \{8\ s, 15\ s, 30\ s, 1\ min, 2\ min, 5\ min\}$ represent the latency of information with a successful server access attempt. In each of the trials, an agent experiences one of the 36 quality of service combinations of information loss and delay.

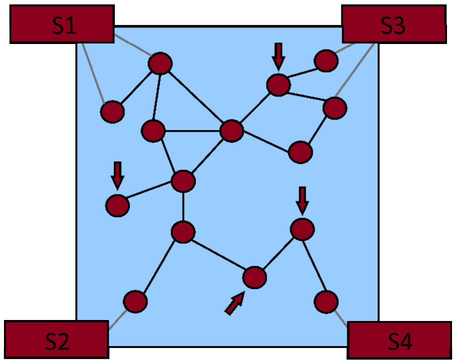

Figure 13. Distributed server scenario.

In figure 14, the correctness versus information loss for three values of information delay is shown. This demonstrates the relative tolerance of decision-making performance with regard to information loss. That is, the sensemaking agents are able to obtain sufficient amount of information in scenarios when there is low to moderate amounts of information loss. As the latency of the communications increases, so does the ability to tolerate losses in the communications. Figure 15 shows the same plot, only as a function of information delay. The relatively graceful degradation of performance as a function of information delay is shown.

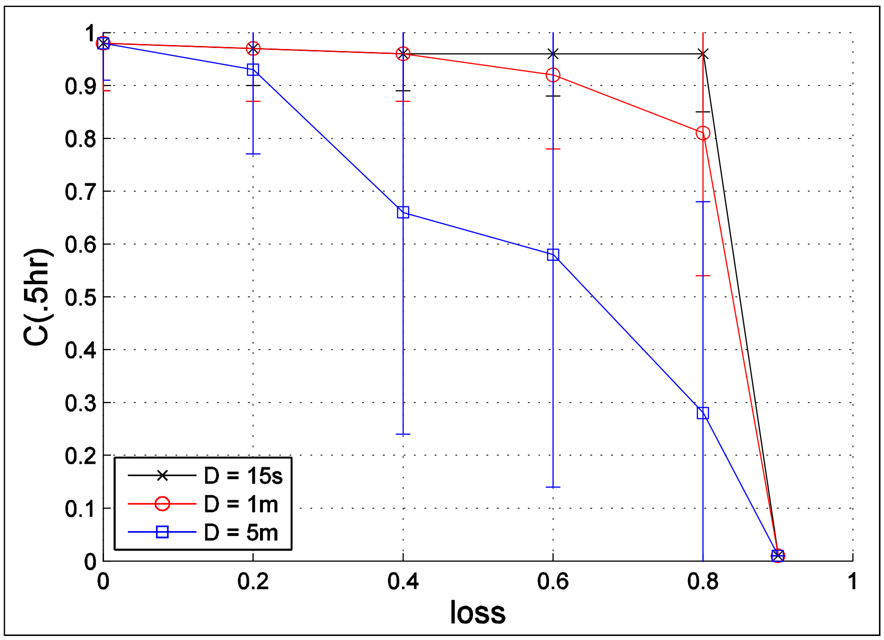

Figure 14. Correctness after 0.5 h vs. information loss for various information delays.

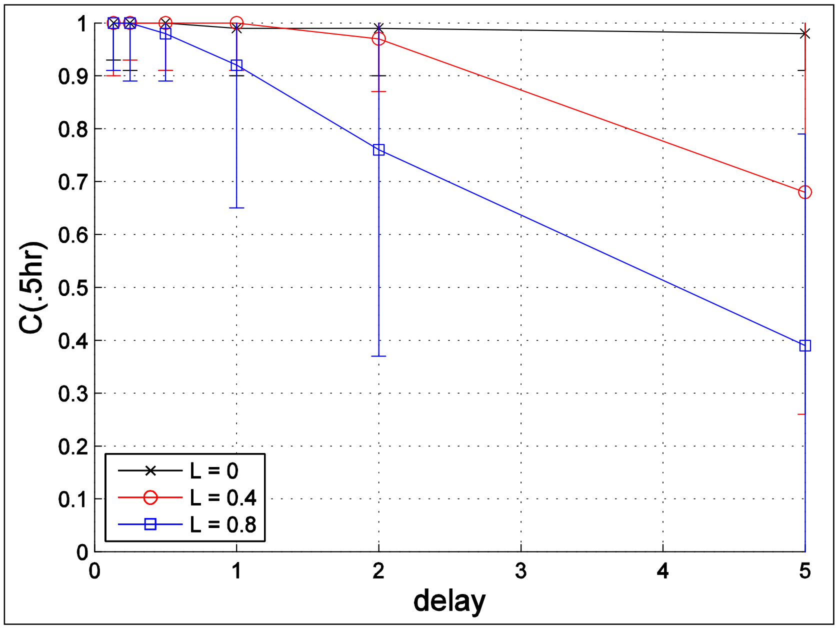

Figure 15. Correctness after 0.5 hr vs. information delay for various information loss.

The results of these experiments indicate that these organizations are more sensitive to delay than loss. A potential explanation for this behavior is the presence of redundancy in communications. If the agents perform an adequate amount of factoid shares in the trials, then there will be a significant amount of duplicate factoids in different nodes within the network. Even with a moderate amount of information loss in the network, the organization can still overcome this with information redundancy. This provides some initial insight into the future design of information sharing protocols within tactical networks. When information loss and information latency are at reasonable levels, the payoff of minimizing information latency is greater than reducing information loss in terms of decision-making performance.

4. Conclusions

4.1 Transitions

Extensions of this research have been submitted as part of a proposal entitled "Distributed Trust Management Schemes and Impact on Network Security" to the Office of the Secretary of Defense's FY11 Cyber Operations Applied Research and Advanced Technology Development research program.

4.2 Continuing Research

This Director's Strategic Initiative (DSI) research effort has led to continuing research in several different tracks:

1. We are currently exploring opportunities to perform these experiments with human subjects. This will enable validation of these results and provide insight into the differences between human and agent behavior in this scenario.
2. The communications in these experiments were represented by the sensemaking agent parameters. By adding more realistic communication models to the experiments, we may be able to obtain results with greater significance. We are currently integrating the ELICIT framework within the wireless emulation laboratory (WEL). This will add real radio models, congestion effects to the communication during the ELICIT experiments.
3. The agents in ELICIT are limited in their ability to consider the dynamics of trust of other participants (human or agent). We are currently considering enhancements to the ELICIT agent to modify their processing or sharing strategies based on trust in other entities or websites. That is, their decisions to accept shared factoids or to share factoids with someone will largely depend on their given history with the other node.

5. References

1. Committee on Network Science for Future Army Applications, National Research Council *Network Science*, The National Academies Press, 2005, ISBN: 0-309-10026-7.
2. Wei, J.; Su, C. eQoS: Provisioning of Client-Perceived End-to-End QoS Guarantees in Web Servers. *IEEE Transactions on Computers* **December 2006**, *55* (12).
3. Small, T.; Haas, Z. Resource and Performance Tradeoffs in Delay Tolerant Wireless Networks. *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking,* August 2005.
4. Cui, S.; Madan, R.; Goldsmith, A.; Lall, S. Energy-Delay tradeoff for data collection. *Proc. of IEEE Intl. Conf. on Communications (ICC),* Seoul, Korea, May, 2005.
5. Famaey, J.; Wauters, T.; De Turck, F.; Dhoedt, B. Towards Efficient Service Placement and Server Selection for Large-Scale Deployments. *IEEE The Fourth Advanced International Conference on Telecommunications*, 2008.
6. Satyanarayanan, M. Fundamental Challenges in Mobile Computing. *Proceedings of the Fifteenth ACM Symposium on Principles of Distributed Computing*, Philadelphia, PA, May 1996.
7. Brehmer, B. Command Without Commanders. *14th International Command and Control Research and Technology Symposium*, June 2009.
8. Fischbach, K.; Gloor, P.; Schoder, D. Analysis of Informal Communication Networks – A Case Study. *Business & Information Systems Engineering* **2009**, *2*.
9. Boyer, H. Simulating ITA Scenarios using Battlefield 2. *ACITA* **Sept. 2008**, *2008*.
10. Mezzacappa, E.; Cooke, G.; Yagrich, K. Network Science and Crowd Behavior Metrics. *27th Army Science Conference*, 2008.
11. Lazer, D.; Friedman, A. The social structure of exploration and exploitation. *Administrative Science Quarterly* 52 **December 2007**, 667–694.
12. Diesner, J.; Frantz, T.; Carley, K. Communication Networks from the Enron Email Corpus, "It's Always About the People. Enron is no Different". *Computational and Mathematical Organization Theory* **2005**, *11*, 201–228.
13. Ruddy, M. The Experimental Laboratory for Investigating Collaboration, Information-sharing and Trust. *Proceedings of the 12th International Command and Control Research and Technology Symposium*, Newport, RI, June 2007.

14. Alberts, D.; Hayes, R. Power to the Edge. *CCRP*, 2003.

15. Ruddy, M. Instantiation of a Sensemaking Agent for use with ELICIT Experimentation. *Proceedings of the 14th International Command and Control Research and Technology Symposium*, Washington, DC, June 2009.

16. Wynn, D.; Ruddy, M.; Nisseny, M. Tailoring Software Agents to Emulate Specific People. *Proceedings of the 15th International Command and Control Research and Technology Symposium*, Washington, DC, June 2010.

17. Chan, K.; Ivanic, N. Connections Between Communications and Social Networks Using ELICIT. *15th International Command and Control Research and Technology Symposium (ICCRTS),* Santa Monica, CA, U.S.A., 22–24 June. 2010

No of. Copies	Organization
1 (PDF ONLY)	ADMNSTR DEFNS TECHL INFO CTR ATTN DTIC OCP 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
1 HC	US ARMY RSRCH LAB ATTN RDRL CIM G T LANDFRIED BLDG 4600 ABERDEEN PROVING GROUND MD 21005-5066
3 HCs	US ARMY RSRCH LAB ATTN IMNE ALC HRR MAIL & RECORDS MGMT ATTN RDRL CIM L TECHL LIB ATTN RDRL CIM P TECHL PUB ADELPHI MD 20783-1197
2 HCs	US ARMY RSRCH LAB ATTN RDRL CIN T K CHAN B RIVERA 2800 POWER MILL ROAD ADELPHI MD 20783

TOTAL:7 (1 ELEC, 6 HCs)

INTENTIONALLY LEFT BLANK.